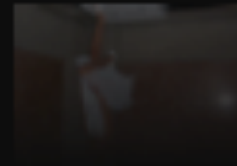
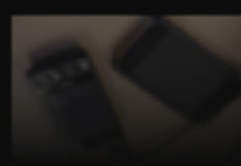
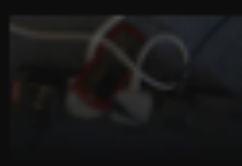
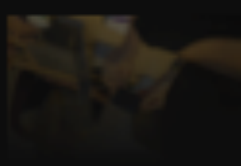
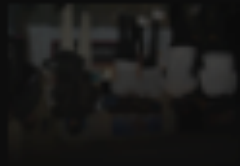
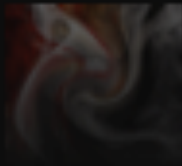
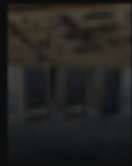
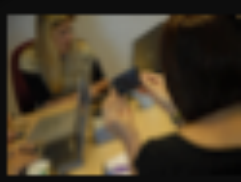
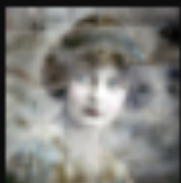
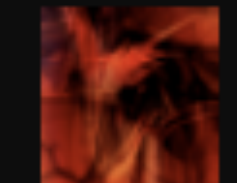
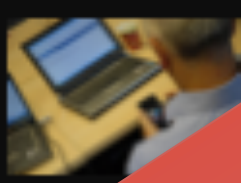
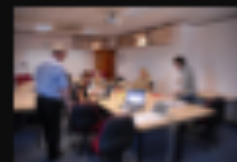
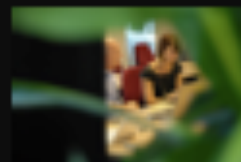
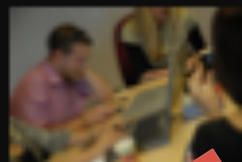
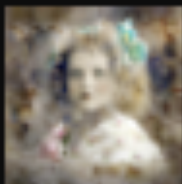


REACTION



REACTION:

Retrieval, Extraction, and
Aggregation Computing Technology
for Integrating and Organizing News

Mário J. Silva

University of Lisbon, Portugal

REACTION

The Problem...

- **Computational journalism**, aka *database journalism*
 - Intensive use of software tools for news research, production and presentation
- What is the **impact in the routines of newsrooms?**
- What **effect** will these tools have on the **quality of news** and the **productivity of journalists?**

REACTION

Challenges

1. Automatic **content analysis**
(documents, news, blogs, micro-blogs, comments)
2. Automatic analysis of
explicit and implicit social networks
3. Design of
rich visualization and interaction interfaces
4. **Case-study** evaluation of developed
computational journalism methodology in a
production setting.
Critical analysis of practical impact on newsroom
quality, efficiency, and economics.

REACTION

Directions

- Automatic Content analysis
 - **Semantic annotation**, involving subjectivity analysis and the **identification of opinions** in context
- Explicit and implicit social networks analysis
 - Entity ranking, **expert finding**
 - Research shingling for detecting and tracking popular passages, **memes**, across news.
- User interface design and analysis
 - Provide information navigation tools and automatic detection of relevant events **to journalists**.

REACTION

Partnership

- **LASIGE, FCUL**
(Mário J. Silva, Paula Carvalho, Francisco Couto)
- **LIACC, FEUP**
(Eugénio de Oliveira, Eduarda M. Rodrigues, Luís Sarmento)
- **CIMJ** (António Granado)
- **Austin: School of Information and Computer Science at Austin**
(Luis Francisco-Revilla, Matthew Lease)
- **PT Comunicações, SAPO**
(Benjamim Júnior, Celso Martinho, Luís Sarmento)
- **Público** (Sérgio B. Gomes)

Research tasks

1. Information Mining
2. Information Discovery
3. Web Community Sensing
4. Tracking Information Flow
5. Interaction and Personalization
6. Query and Visualization
7. Computational Newsroom

REACTION

Information Mining

- Development of robust **linguistic resources** to process different types and genres of texts
 - knowledge resources about media personalities: **recognizing and resolving references to named-entities;**
 - sentiment lexicons and grammars: detecting the **polarity of opinions about relevant personalities**
 - annotated corpora: training different text classifiers and evaluating classification procedures

Information Discovery

Relationship extraction techniques to support
information discovery in journalists' activities

- **Entity Ranking:** finding the relevant entities for a given topic
- **Entity Distillation:** finding relevant resources for a given entity
- **Attribute Selection:** finding a list of key aspects to compare and differentiate a given set of entities

REACTION

Web Community Sensing

- Modeling the credibility and authority of news sources and **opinion makers in social networks**
- Identifying **influential individuals and experts** on a given news topic
- Monitoring the **community reaction to news** stories and the polarity of opinions

REACTION

Tracking Information Flow

- Identifying **originating source** of new ideas and information
- Understand **evolutionary development** of ideas through their **iterative retelling and revision** over time and across sources
 - detecting cases and patterns of re-use (e.g. via “**memes**” or larger units of similar text) and information flow for source identification and novelty detection.

REACTION

Interaction and Personalization

- Determining which **interaction and personalization** mechanisms are best suited to:
 - Significantly enhance the user experience
 - Provide the news site with useful, tacit feedback about its readers' needs
- Investigating interactive **news interfaces** that support both automatic and manual personalization for readers

REACTION

Query and Visualization

- Development of tools for querying extracted information and **visualizing annotated documents** and datasets
- Continuous scanning of the social web, news sources and various kinds of data streams
 - Sapo already scans and processes many of these streams, in particular the news media

Computational Newsroom

- **Environment** where the **new tools and resources** developed in the project, together with other software will be accessible
- Will **use tools and collect data** for case studies to be evaluated
 - observation and structured interviewing of the journalists in contact with the developed tools.
- The research will try to **contextualize the changing nature of media work**

REACTION

More details

- Starts October 1st, 3 years
- <http://xldb.fc.ul.pt/wiki/Reaction>

UT Austin | Portugal

INTERNATIONAL COLLABORATORY FOR EMERGING TECHNOLOGIES, CoLab

REACTION